DOCUMENT RESUME

ED 244 731 PS 014 291

AUTHOR St. Pierre, Robert G.
TITLE Three Ways of Learning More from Follow Through: Secondary Analysis of Extant Data, Compilation and Analysis of Follow-Up Data, and Completely New Studies.
INSTITUTION Abt Associates, Inc., Cambridge, Mass.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE 15 Jan 81
NOTE 42p.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Data; Early Childhood Education; *Followup Studies; *Research Needs; *Research Proposals
IDENTIFIERS *Project Follow Through; *Secondary Analysis

ABSTRACT

This discussion argues that additional information can be obtained from the Follow Through program in its current framework. Three ways of generating new knowledge from Follow Through are suggested: (1) secondary analysis of extant data, (2) compilation and analysis of followup data, and (3) completely new studies. Two ways of minimizing costs through the use of existing data on children who have already completed Follow Through programs are also mentioned. In addition, the discussion reviews and comments on some of the suggestions for Follow Through research studies that have been proposed in the 4 years since completion of the national evaluation. (RH)

Three Ways of Learning More From Follow Through: Secondary Analysis of Extant Data, Compilation and Analysis of Follow-Up Data, and Completely New Studies

Robert G. St.Pierre

Abt Associates Inc.
55 Wheeler Street
Cambridge, Massachusetts 02138
617/492-7100

January 15, 1981

Three Ways of Learning More From Follow Through: Secondary Analysis of Extant Data, Compilation and Analysis of Follow-Up Data, and Completely New Studies

This paper has been prepared for the National Institute of Education as part of their agreement with the Office of Elementary and Secondary Education to participate in new Follow Through (FT) research and pilot project activities. The paper's objective is to address three questions related to FT's potential for generating new knowledge:

- What more can we learn from FT through secondary analysis of extant data? What FT-related data bases exist at the national, sponsor, and site levels?
- What can we learn from compiling and analyzing follow-up data on children who participated in FT?
- What can we learn from new studies of sponsors and sites in FT's current framework?

Implicit in the title of this paper is the conviction that there is more to be learned from FT. This belief stems from FT's history as well as from the fact that FT now has a dual focus. New regulations mandate that 80 percent of FT's funds be used for the provision of service while the remaining 20 percent are to be used for knowledge production (Wholey 1979). The regulations thus reinforce the vision of FT as a federally funded education program containing a

built-in research component. There are few enough programs with a mandate to do educational research and knowledge production, and enough of a need for research, so that by any definition there is a substantial amount to be learned from FT.

The national FT evaluation (Stebbins, St.Pierre, Proper, Anderson and Cerva 1977) showed that in terms of certain outcomes, FT children in some models performed better than comparison children, some performed about the same, and some did worse. These results are not particularly upsetting if we see FT as a research program. From this viewpoint, the national evaluation was not an evaluation of FT (which would have assessed the adequacy of FT as a vehicle for providing a testing ground for the examination of the effects of model sponsors), but was a part of FT.

Regardless of the outcomes of the national evaluation, FT continues to have worth as a service and knowledge production program designed to develop, test, and disseminate alternative models of education. This statement ignores the debate over claims of model effectiveness and is based on the opinion that FT addresses an important problem. It is the largest and most sophisticated vehicle that the federal government has for testing the effects of alternative educational models; it functions to provide funds for curriculum research and development under field conditions; and it has already generated a wealth of knowledge about primary school education, the implementation of externally-developed programs in the public schools, and evaluation methods for field-based programs. Because of this history, and the prospect for continued research, development, testing, and dissemination of educational practices, FT has probably come closer to being a "reform as an experiment" (Campbell 1969) and to Campbell's (1971) vision of the "experimenting society" than any other federal program.

There certainly is room for disagreement with this opinion. In fact, there are two widely held views that lead to very different conclusions about whether or not we can learn more from FT. One of these is held by many FT practitioners, administrators and Congressional supporters--that FT is a service program and only a service program. In this view resources allocated to FT research are resources not allocated to service and by definition are ill-spent. Clearly, a person holding this view does not see FT as a prime vehicle for generating new knowledge. The second opposing view, held by many in the current administration, is that FT is simply an expensive version of Title I and should be phased out.

The fact that persons with these three major views (FT as a program capable of supporting both service and important research, FT solely as a service program, and FT as a program to be phased out) have to interact in the process of planning FT research has lead to great difficulty in organizing, planning, and carrying out any sort of research. This point will be addressed again later.

This paper is written under the assumption that the real question for persons holding the first viewpoint articulated above is not whether we can learn more from FT in its current framework, but how much in the way of scarce resources should be invested in gaining what kinds of knowledge. The paper discusses two ways of minimizing such an investment through the use of existing data on children who have already completed the FT program. In addition, it reviews and comments on some of the suggestions for new FT research studies that have been proposed in the four years since completion of the national evaluation.

1.0 What More Can We Learn from FT Through Secondary Analysis of Extant Data?

1.1 Motivation for Studies Based on Extant Data

There are important questions to be answered about and by FT. Further, relevant data sets exist or can be constructed from school records to address some of these questions. While the constraint of relying on accessible, existing data limits the range of questions that can be answered, the advantages of performing secondary analysis (Cook 1974) are considerable. Most important are that studies based on extant data will be less expensive and more timely than those involving substantial new data collection.

Secondary analysis requires minimal resources since costs associated with designing the original evaluation, developing instruments, collecting the data, and building the data base have all been met by others (Boruch and Reis 1980). The secondary analyst must bear only the costs of data access, analysis, and report writing. Further, because it shortcuts the startup, planning, data collection, and data base construction periods of an evaluation, secondary analysis of extant data can yield information in a fraction of the time that it would take to mount a new study.

The major drawback to secondary analysis is that the questions that can be addressed are limited by the nature of the extant data. It is often the case that the secondary analyst has different or more detailed questions than were asked in the primary evaluation. Variables that are key to the secondary analyst may be missing from the data base. It should also be recognized that for some questions (e.g. those that are most easily addressed through case studies or other qualitative methods) secondary analysis may be infeasible or less cost effective than the collection of completely new data. For many such questions the data simply do not exist, or if they do, the circumstances of their collection are so unique that the data have little applicability to other

situations. However, assuming data exist to support research on important questions, secondary analysis is an efficient and timely mechanism.

1.2 What Extant Data are Available for FT?

Many data sets related to various aspects of FT are available including the national evaluation data base, several sponsor and site maintained data bases, and data sets built through other federally funded FT research projects.

1.2.1 National Evaluation Data Base

Perhaps the most obvious, though not necessarily the best, extant resource for learning more from FT is the data base resulting from the ten-year national FT evaluation. Data on over 20 FT sponsors that implemented their educational models in more than 200 sites nationwide are contained in this massive data set. The data base is organized into four cohorts which comprise FT and NFT children from all sponsors and sites. Within each cohort the data base is hierarchically organized by sponsor, site within sponsor, school within site, class within school, and child within class. Contained within these sets are various types of data including child level demographic data, child level test scores, child level parent interview data, and classroom level teacher and teacher aide data. For some children, only demographic data are present; for many, the other types of data are present for up to four years (grades K-3). Tables 1, 2 and 3 are adapted from Goodrich and St.Pierre (1979) and give some details on the SRI data base.

Middleton and Durgin (1978) have documented the national evaluation data base as constructed by SRI International. However, the national evaluation data set is no longer maintained by SRI. A copy has been submitted to the National Archives and to this writer's knowledge at least one attempt to

Table 1

Demographic Data on the SRI Data Base*

Variable	Cohort	I		II			III		IV
	Stream	EK	EF	EK		EF	EK	EF	EK
	Grade	K,1,2,3	1,2,3	K,1,2	3	1,2,3	K,1,2,3	1,2,3	K,1,2
Project		x	x	x	x	x	x	x	x
School		x	x	x	x	x	x	x	x
FT/NFT designator		x	x	x	x	x	x	x	x
Grade		x	x	x	x	x			
Child birthdate		x	x	x	x	x	x	x	x
Sex		x	x	x	x	x	x	x	x
Ethnicity		x	x	x	x	x	x	x	x
First language spoken in the home		x	x	x	x	x	x	x	x
Second language spoken in the home		x	x	x	x	x	x	x	x
Head Start flag		x	x	x	x	x			
Head Start equivalent flag					x				
Months Head Start or equivalent experience		x	x	x	x	x			
Months Head Start experience							x	x	
Months equivalent experience							x	x	
Days absent during year		x	x	x	x	x	x	x	x
Child eligible for FT services					x		x	x	x
Classroom service available					x		x	x	x
Lunch service available					x		x	x	x
Medical, dental service available					x		x	x	x
Other service available					x		x	x	x
Date child entered class					x				
Date child left class					x				
Years of data available on child							x	x	x
Parent interview history							x	x	x
Child test history							x	x	
Child FT/NFT by year							x	x	x
Months in FT prior to this year							x	x	x
Child entering grade									x

Table 2

Test Data on the SRI Data Base[1]

MEASURE	COHORT	I										II										III												IV			
	STREAM	EK						EF				EK					EF					EK					EF						EK			
	GRADE	K	1		2		3	1	2		3	K	1	2		3	1	2		3		K	1		2	3	1	2		3		K	1		2		
	TEST POINT	F69 S70	F70	S71	F71	F72	S73	F69 S70	F70	S71	S72	F70 S71	S72	F72	S73	S74	F70	S71	S72	F72	S73	F71	S72	F72	S73	S74	S75	F71	S72	F72	S73	S74	F72	S73	S74	S75
WRAT		x	x	x				x	x	x		x					x	x				x	x					x	x				x			
PSI		x						x				x					x	x				x						x								
LEE-CLARK		x						x				x					x	x				x						x								
EC1		x						x				x					x	x																		
FACES		x	x	x				x	x	x		x					x	x				x						x								
SOCIAL INTEGRATION ITEMS		x	x	x				x	x	x		x					x	x																		
MRT			x	x					x	x								x																		
SPONSOR ITEMS (VERBAL & QUANT.)			x	x													x	x																		
SPONSOR ITEMS (MISC.)			x	x					x	x																										
MAT PRIMER																						x	x										x			
MAT PRIMARY I									x				x	x											x			x	x					x		
MAT PRIMARY II					x	x				x					x				x	x						x				x						x
MAT ELEMENTARY							x				x					x				x							x				x					
RAVEN'S					x		x				x					x			x							x				x					
COOPERSMITH					x		x				x					x			x							x				x					
IARS					x		x				x					x			x							x				x					
GUMPOOKIES													x										x					x					x			
LOCUS OF CONTROL													x										x					x								
PPVT																						x	x				x	x				x			

Table 3

Counts of Children with Varying Degrees of FT Exposure and Test Data

SPONSOR	COHORT III-EK				COHORT II-EK				COHORT III-EF				COHORT II-EF			
		In FT All Four Years				In FT All Four Years				In FT All Four Years				In FT All Four Years		
	In FT Some-time	No Test Require-ment	Post-tested	Pre- and Post-tested	In FT Some-time	No Test Require-ment	Post-tested	Pre- and Post-tested	In FT Some-time	No Test Require-ment	Post-tested	Pre- and Post-tested	In FT Some-time	No Test Require-ment	Post-tested	Pre- and Post-tested
Self-Sponsored*	644	354			699	364	208	179	223	180			301	282		
Far West Labs	693	367	230	215	695	303	308	285	550	495	294	293	561	505		
Arizona	966	639	341	317	782	594	354	307	745	561	416	410	769	689		
Bank Street	668	432	270	250	581	349	266	234	441	354			397	377	81	71
Georgia	107	101			118	109			419	301	292	283	328	305		
Oregon	958	636	251	241	772	565	281	250	771	583	278	236	822	749	225	136
Kansas	1097	802	621	559	806	679	369	341	59	49			223	200		
High/Scope	361	222	147	131	218	166	114	108	215	210			237	222	125	45
North Carolina	507	335	241	230	570	421	269	239	561	447			407	342	191	174
EDC	446	230	222	201	300	172	132	92	285	231			278	264	239	210
Pittsburgh	342	261	254	235	172	141			80	67			80	78		
Fordham	274	129	123	116	239	205	54	49								
SEDL	703	404	393	357	575	329	220	183	107	104			227	218		
Hampton	60								43	31			54	31		
Northeastern Ill.	177	73	62	61	185	169										
Georgia State									295	242			173	148		
Clark College	43	34	33	32	40	34	34	31	155	137			78	71		
Calif. Dept of Ed**	390	185	157	153	432	280	195	187								
North Dakota	156	78	55	55												
TOTAL	8592	5202	3400	3153	7264	4960	2804	2485	4949	3992	1280	1222	4935	4481	861	636

*FT Sponsors are referred to by their institutional affiliation.

**No longer an FT sponsor.

obtain the data through that route was successful. Abt Associates Inc. (AAI) maintains the analytic files used in the national evaluation on an ad hoc basis and has distributed copies to researchers upon request.

1.2.2 FT Sponsor and Site Supported Data Sets

The national evaluation data base includes many sites and sponsors with a common, but restricted, set of measures. It has breadth but lacks depth. On the other hand, internal evaluations done by several FT sponsors (c.f. Hodges and Sheehan 1978) complement the national effort by adding depth in selected instances. Some sponsors developed their own instruments, others used standard tests not employed in the national evaluation. These efforts were idiosyncratic but contain information that is sometimes more relevant to the sponsors' objectives than the national data.

While studying the utility of sponsor and site maintained data sets for supporting research on the long-term effects of FT, Goodrich and St.Pierre (1979) surveyed all FT sponsors and several FT sites as to the existance and nature of relevant data sets. Table 4 summarizes the results of this survey for some of the sponsors/sites having data sets most amenable to secondary analysis. Several of these data sets could support research on the effects of FT during and subsequent to the FT years.*

Sponsor/site data sets vary tremendously in quality. The best (i.e. some of those shown in Table 4) are characterized by several years of longitudinal data on FT and NFT children, well matched comparison groups, pretest and posttest data and many measures. Some contain measures specifically designed to tap

*This survey was current as of early 1979. Since that time some sponsors have undertaken significant data base development and any thorough examination of sponsor data should include an update of the current status of these data sets.

Table 4

Summary of Selected Sponsor/Site Data Bases

Data Base Maintained By	Sites/Sponsors Included in Data Base	Grades	Measures	Comparison Groups	Estimated Longitudinal Sample Size
University of Arizona/ Tucson Early Education Model	Wichita, KS	1-6	Achievement tests, productive language test	Randomized across treatment/control, other nonequivalent comparisons	40 (of initial 100) per cohort for several cohorts
Fordham University/ Interdependent Learning Model	Atlanta, GA New York City, NY	1-6	Achievement tests, problem solving, phonic skills, many noncognitive measures	Well-matched local comparison groups in Atlanta, fair to poor in New York City	1,000's in Atlanta, 100's in New York City
Georgia State University/ Parent Supported Diagnostic Model	Natchitoches Parish, LA	1-6	Achievement tests, attendance	Well-matched local comparison groups for three of six schools	Low 100's across all cohorts
University Of Kansas/ Behavior Analysis Approach	New York City, NY Philadelphia, PA Portageville, MO Trenton, NJ Kansas City, MO Louisville, KY Waukegan, IL Meridian, IL N. Cheyenne, MT Hopi, AZ	K-5	Achievement tests, attendance, consumer satisfaction	Well-matched local comparison groups in urban sites, poor or no comparisons in other sites	100's per cohort for several cohorts
University of Oregon/ Direct Instruction Model	Flippin, AK E. St.Louis, IL Smithville, TN Uvalde, TX Dayton, OH Tupelo, MS Flint, MI	K-6	Achievement tests, attendance	Well-matched local comparison groups in most sites	100's per cohort for two cohorts
Southwest Educational Development Laboratory/ Bilingual/Bicultural Model	Cutler-Orosi, CA Los Angeles, CA San Diego, TX St. Martin Parish, LA Tulare, CA	K-6	Achievement tests, Spanish reading	Adequately matched local comparison groups	75-150 per site per cohort
Philadelphia Public Schools	Bank Street EDC Kansas North Carolina Phila. Process Parent Implem. SEDL	1-8	Achievement tests, attendance, info. on competing treatments	Cross-sponsor comparisons, well-matched local comparison groups up to 1975	100's per sponsor per cohort

features of the sponsor's educational model. For example, the Fordham data base contains measures of problem solving as well as many other noncognitive measures, while the Oregon and Kansas data sets are strong in terms of academic achievement measures.

Some of the data sets are very strong in terms of design. For example, the data set supported by an Arizona site (the Wichita Public Schools) contains data from a classical experiment where Head Start-eligible children were assigned randomly to FT and NFT groups across several cohorts.

With a few notable exceptions (e.g. Philadelphia, Fordham) most sponsor and site data bases are not readily transportable. They tend to exist in bits and pieces and would require a modest effort to pull together. Even some of the best data sets in terms of measures and design (e.g. Kansas, Oregon) exist in cross-sectional files that would require merging prior to longitudinal analysis.

1.2.3 Other Research Based Data Sets

While the national evaluation and the FT sponsors are the primary sources of extant data, some ongoing research studies will provide important data sets. First, System Development Corporation is conducting a study of parental involvement in four federal programs (FT, Title I, ESAA, and Bilingual Education). In the course of this study (which is not yet completed) SDC researchers collected two types of data. A national survey was conducted to determine the nature and level of parent involvement activities in FT (as well as in the other federal programs noted above). The data base from this survey will be available in the near future. The second type of data collected was indepth case study data on a sample of 16 FT sites that were determined through the national survey to be active in terms of parental involvement. The data base for this study is not yet complete, but it will likely consist of hard copy case study materials.

Second, the Network for Innovative Schools has conducted a study of FT Resource Centers (Shive, Meleen, Harris, Vaughn and Grogan 1980). Using case study methods Network researchers developed individual profiles of about 25 resource centers which were then used to prepare a report on the resource centers as a whole. The profiles comprise the study's data base and they could be made available with some effort devoted to preserving the anonymity of the sites.

Finally, the National Urban League is beginning a multi-year assessment of the social implications of FT. This study will examine the impact of FT in 12 urban cities. It will look for indirect effects of FT such as increased community involvement in schooling as well as improved health and social services. The study is just starting and data will not be available for many months.

1.3 Examples of Questions that Can be Addressed by Extant Data

This discussion will concentrate on relevant study questions that can be addressed by the data sets noted above. It focuses on questions related to the immediate effects of FT on children, parents and teachers, i.e. the effects of FT on program participants while they are in the program. It should be understood that this discussion (and other discussions of research questions) is not intended to provide a census of all possible study questions. Rather, examples which illustrate the range of potential research questions will be identified.

1.3.1 Questions Relevant to the National Evaluation Data Base

Several questions about the immediate effects of FT can be addressed by using the national evaluation data. There has been limited analysis and even more limited reanalysis of the national evaluation data. The data base is huge and rich, containing multiple measures on multiple cohorts of children, parents

and teachers from grades K-3. In no way were the AAI researchers able to fully explore the depths of this data set. Examples of questions that have thus far gone unaddressed or have not been addressed fully include:

- What is the relationship between amount of FT exposure and achievement? How does this relationship vary by model?
- What effect does FT have on school attendance? How does this relationship vary by model?

The national evaluation concentrated on the estimation of child effects rather than on a thorough analysis of data from FT parents and/or teachers. Noting that these data were not fully explored in the national evaluation, Haney and Pennington (1978) conducted a descriptive analysis in order to learn more about FT as a comprehensive service program, to better understand FT in terms of social action and parent involvement, and to help interpret the results of the national evaluation. Questions that could be explored further in this area include:

- What has been the role of FT teachers and teacher aides? How does the role differ by model?

It is also important for researchers and metaevaluators to reanalyze the data set used by Stebbins et al. to see if the national evaluation findings and conclusions hold up under close scrutiny. House, Glass, Mclean and Walker (1978) authored the most frequently cited critique of the national evaluation. Yet, in spite of their objections to the analysis strategy used in the national evaluation, House et al. did not reanalyze the raw child level data, but chose instead to accept AAI's site level results and based their "reanalysis" on the aggregate site level data presented in various national evaluation reports (Stebbins et al. 1977; Bock, Stebbins and Proper 1977; associated appendices). Thus, regardless of their complaints about the analysis strategy, House et al. by definition accepted the basic analysis techniques employed by the original

evaluators. The same criticism can be leveled at other FT metaevaluators (e.g. Bereiter and Kurland 1978).

Thus, there is work to be done with the national evaluation data base simply in terms of checking the AAI analyses for reliability:

- How reliable are the national evaluation results? Do they stand up under the scrutiny of secondary analysis?
- Do different researchers using different analysis methods making different assumptions arrive at different conclusions about FT? About individual FT models?

To this writer's knowledge, Camilli's (1980) doctoral thesis presents the only child level reanalysis of the national evaluation data that has been done to date. Among other analyses, Camilli tried several solutions to the problem of covariance adjustment, explored the generalizability of the national evaluation model and instrument categories, and examined the relationships of parent and teacher variables to outcome gains. According to Camilli, his "...reanalysis does not produce shocking discrepancies..." with the national evaluation findings.

1.3.2 Questions Relevant to Sponsor and Site Data Bases

Other sources of information, in addition to the national evaluation, are relevant to FT. These include, for example, sponsor and site data sets. Many of the questions about model effectiveness addressed in the national evaluation can be addressed using sponsor data. However, given the diversity in measures and designs, model comparisons or "horse races" would not be possible except at the most global level:

- What are the effects of participation in a given FT model on achievement? On affect? On attendance?
- What can be said about causal relationships among these outcomes?

The sponsor data sets also contain unique information such as Kansas' data on consumer satisfaction with the program. In this case, an annual survey of program children, parents, teachers and administrators is used to gather data on perceptions about the program, level of satisfaction, and suggestions for improvement.

- How satisfied are FT consumers (parents, teachers, children)?
- What is the relationship between parent satisfaction and child achievement? And attendance?

Another interesting and neglected area of work could concentrate on agreements and disagreements about FT when viewed from the perspectives of different data sets. Clearly, evidence apart from the national evaluation should be integrated into an overall picture of FT:

- Do different FT data sets (national evaluation, sponsor, site) lead to different conclusions about the success of FT models? If so, why?

Finally, no data were ever collected in the national evaluation as to the degree to which the FT models were implemented. Questions as to the conformity of the treatments to sponsors' specifications or even as to the existance of the treatments were never addressed systematically. Some sponsors have since invested substantial resources in documenting the implementation of their programs and some of their implementation data sets are worthy of analysis.

- How have FT sponsors gone about studying the implementation of their programs? Do their methods have implications for the study of implementation in other programs?
- Do the data that FT sponsors have collected on program implementation allow distinctions to be made among FT models? Do the data accurately describe what sponsors think should be happening in their classrooms?
- By their own analysis are FT sponsors able to implement their models? Does external reanalysis confirm their conclusions?

2.0 What Can We Learn from Compiling and Analyzing Follow-Up Data on FT Participants?

The possibility of FT models having delayed effects (sleeper effects, long-term effects) has been raised at several points over the past decade. The original purpose of FT was to improve the life chances of children from low income families. From this perspective it makes sense to devote attention to post-FT effects, especially considering that concern over the apparent fade-out of Head Start effects (Wolff and Stein 1966) was the primary impetus behind the establishment of FT. However, as Goodrich and St.Pierre (1979) note "Studies of post-treatment effects have been seen as interesting but not timely..." This section of the paper addresses the question of studying delayed FT effects.

2.1 Motivation for Conducting Follow-Up Studies

In spite of relevant planning work by Goodrich and St.Pierre (1979), Garet (1977), Madaus and Marx (1973) and a proposal by the School District of Philadelphia (1980), studies of delayed FT effects have not been funded at the federal level. Several reasons for studying post-FT effects are given in these sources and are summarized here. First, FT was initiated because the achievement gains obtained by Head Start children appeared to dissipate in the early school years. The same issue is of concern to FT--if achievement gains are produced by FT, do they dissipate, or are they sustained after completion of the program?

Second, many FT sponsors try to impact childrens' affect, conceptual ability, cognitive style, other non-achievement areas and parental behavior. These sponsors believe that achievement gains will be obtained indirectly through impacts on mediating variables such as those listed above. In this case, effects on student achievement are not expected to occur quickly. In order to evaluate these sponsors' models it is necessary to examine achievement

gains subsequent to program termination. The question is, do FT models that aim for achievement gains by first impacting mediating variables produce delayed or "sleeper" effects?

A third motivation for studies of post-FT effects is that FT projects, by design or as a byproduct, may affect parental and community opinions. In general, parents support FT and have been a potent force in maintaining the program in the face of efforts to phase it out. The question which might arise is, what are parents', teachers', and students' opinions of FT once students have made the transition to "regular" public school classes?

Fourth, FT's overall objective has been to improve the "life chances" of children from low income families. College attendance, future earnings, social mobility and other post-school variables are all important indicators of success in life but are not available measures in any existing post-FT data base. Life chances studies could concentrate on earlier, more accessible indicators of success such as school attendance, special education placement, grade retention, dropout rate, grades, course selection, and discipline records.

Although post-FT studies have not been funded at the federal level, interest has been high among sponsors where debates about "sleeper" effects and fade-out have taken place regularly. Studies of varying degrees of sophistication have been conducted by the following sponsors: Arizona (Riley 1978; Cloud, Rentfrow, and Hildebrandt 1979), Bank Street (Seitz, Apfel, and Efron 1977), Far West Labs (Edwards and Bridewell 1979), Kansas (assorted reports), and Oregon (Weber and Fuhrmann 1978; Becker and Englemann 1978).

2.2 What Follow-Up Data are Available?

It is clear that some sponsors have relevant data. Based on their survey of sponsor data sets, Goodrich and St.Pierre (1979) recommended six data sets

as being appropriate for producing new knowledge about post-FT effects. Three of the data bases, those supported by Philadelphia, Fordham and Oregon, afford especially good opportunities since they meet fairly stringent methodological standards (ample sample size, several sites, a wide range of measures, good local comparison groups). The other three data sets, maintained by Arizona, Georgia State and SEDL, would support smaller studies in single sites that are somewhat weaker methodologically. It should be noted that while other sponsor data sets were not recommended for study by Goodrich and St.Pierre, several FT sponsors have upgraded their data management systems during the past year or two, and new data bases may be available.

2.3 Examples of Questions that Can Be Addressed by Follow-Up Studies

Table 5 presents study questions of particular interest in this area that could be addressed by the best of the sponsor and site supported data bases. In spite of the high quality of some sponsor/site data bases, they tend to concentrate on paper and pencil outcome measures, and are relatively weak on "life chances" behavioral variables (e.g. grade retention, special education placement, attendance). Further, none contain follow-up data on the perceptions of children, parents and teachers about FT or data on parents' or teachers' educational/employment status.

There is an important precedent for collecting long-term follow-up data on the school behavior of FT children. The Developmental Continuity Consortium (Lazar, Hubbell, Murray, Rosche, and Royce 1977) analyzed follow-up data that were collected from children aged nine to 19 who participated in 14 different experimental infant and preschool programs. Their study shows positive effects of Head Start on special education assignment and retention in grade (program children were assigned to special education classes and retained in grade less

Table 5*

Examples of Research Questions that Can be Addressed by Selected Sponsor/Site Data Bases

RESEARCH QUESTIONS	SELECTED DATA BASES					
	Phila-delphia	Fordham	Oregon	Arizona	Georgia State	SEDL
• What are the causal relationships among academic achievement, IQ, process measures (amount of FT exposure, attendance, etc.), and affective measures?	X	X	X			
• Do achievement effects associated with participation in FT increase, decrease, or stabilize after completion of the program?						
- If positive effects on academic achievement are produced by FT models at the end of grade three, do they fade-out, or are they sustained in the post-FT years?	X	X	X		X	
- Do FT models that aim for achievement gains by first changing children's affect, conceptual ability, parental behavior, and so on, produce "sleeper effects" on achievement measures?	X					
- How does post-FT academic performance relate to previous FT performance and initial status (e.g. socioeconomic status, pretest, demographic data)? Are gains sustained in the post-FT years?	X	X		X	X	X
• What are the relative importances of the current academic year and the child's previous academic history as determinants of academic success?	X	X	X			
• What is the causal relationship between academic performance and school attendance? In FT? After FT?	X	X	X			
• What is the relationship between post-FT academic performance measures and length of FT exposure?	X	X		X		
• Is it possible to predict post-FT patterns (fade-out, sleeper effects, sustained growth, etc.) on different measures in terms of the FT curriculum?	X		X			
• What differences in post-FT performance (achievement, grades, attendance) exist among different models at the same site?	X					

*Adapted from Goodrich and St.Pierre (1979)

often than comparison children). They also found evidence for fade-out of IQ gains over a three-year period.

While the findings of this study have been controversial, Cloud et al. (1979) investigated the later school success of children enrolled in the FT program in Lakewood, New Jersey and were able to replicate the findings of Lazar et al. Children who participated in FT were found in high school to have significantly lower rates of grade retention and special education placement than comparison groups composed of older siblings. The groups did not differ significantly on variables such as dropout rate, educational achievement, or educational aspirations.

Such work suggests strongly that further research is needed in this area. Discussions with FT sponsors and sites (e.g. Philadelphia) lead to the conclusion that in some instances it would be possible to obtain child level life chances variables from school districts records and append them to existing sponsor/site data sets. If augmented in this fashion some of the sponsor/site data sets would allow researchers to address questions such as:

- What is the long-term impact of a given FT model on life chances variables?
- Do the results of such studies replicate those of Lazar et al. and Cloud et al.?

3.0 What Can We Learn from New Studies of Sponsors and Sites in FT's Current Framework?

The viewpoint set forth earlier, that there is more to be learned from FT, holds as much for completely new studies as for work based on extant data. However, a review of the use of FT research funds since completion of the national evaluation in 1977 shows that it may be easier to generate than to implement ideas for new research.

Table 6 summarizes the work that has been funded with FT research monies from fiscal year 1977 through fiscal year 1980--more or less since completion of the national evaluation. Inspection of this table shows that in the past few years there has been a problem spending FT's research funds. This is due to factors such as the fact that the administration has been trying to phase out the program for years. Instability caused by uncertainty about continuation of funding has lead to great problems in planning a coherent research agenda. In addition, the Division of Follow Through and the Office of Program Evaluation have consistently had conflicting opinions as to the most appropriate research and evaluation activities that should be undertaken. Finally, the conduct of FT research has been complicated by external factors such as the Congressionally imposed cap on spending for consultant services. In 1980 this cap forced the Office of Program Evaluation to return to the U. S. Treasury $2 million that would otherwise have been spent on FT research.

In spite of these structural and organizational constraints there is no dearth of ideas for new FT research, and a good deal of work was contracted between 1977 and 1980. Consideration of the substance of the studies listed in Table 6 reveals that it is possible to discern at least three broad thrusts in

Table 6

FT Research Expenditures (1977-1980)*

Fiscal Year	Contractor	Contract Amount	Description of Contract
1977	AAI	$52,512	Completion of national evaluation
	RMC	35,762	Completion of national cost analysis
	SRI	53,685	Maintenance of national data base
	Huron	21,939	Planning new studies
	Far West Labs	4,500	Paper on FT implementation
1978	AAI	471,511	Planning for later effects studies; search for new models
	SRI	45,490	Maintenance of national data base
	Network	597,618	Multi-program study of dissemination strategies
	SDC	350,000	Multi-program study of parental involvement
	Huron	17,268	Parent/teacher reanalysis
	Rehab	50,769	Study of FT supplemental training
	High/Scope & Kansas	79,278	Study of potential for a joint FT model
	Oregon	60,052	Study of implementation in a new site
	Georgia	53,682	Planning paper for new FT research
	Georgia State	45,352	Planning paper for new FT research
	Kansas	70,739	Planning paper for new FT research
	North Carolina	48,864	Planning paper for new FT research
	Oregon	49,799	Planning paper for new FT research
	SEDL	48,592	Planning paper for new FT research
1979	SDC	450,000	Multi-program study of parental involvement
	SRI	14,825	Archive national data base
	Philadelphia	9,999	Build data base for delayed effects study
	Rehab	12,730	Study of FT supplemental training
	AMS	300,000	Develop performance indicators and systems for monitoring performance of FT sites
	Oregon	143,418	Study of implementation in a new site
	North Carolina	28,911	Data analysis
	Boon-Young	295,321	Design of studies to be conducted by pilot projects including extension of FT beyond grade 3, effects of alternative levels of FT services, capabilities of LEAs
1980	National Urban League	987,279	Assessment of social implications of FT in urban cities

TOTAL BY YEAR

Year	Total Spent**	Total Allocated	Allocated-Spent
1977	$168,398	$1,000,000	$831,602
1978	1,989,014	2,000,000	10,986
1979	1,255,204	1,273,000	17,796***
1980	987,279	3,000,000	2,012,721

* The data for this table were supplied by Jerry Burns of ED/OPE.

** This reflects the total contracted funds. Other small amounts of FT research funds were spent for items such as conferences and travel.

*** $2 million returned to the U. S. Treasury because of Congressionally imposed cap on consultant services.

recent FT research: (1) planning new FT research; (2) improving FT management and implementation; and (3) new FT research (see Table 7).

Since this paper focuses on information to be learned from new FT studies, we will consider further selected planning studies shown in Table 7 as well as other less formal efforts. In order to benefit from the thinking of others, the next section presents a review of several proposals for new FT research.

3.1 Review of Research Proposals from FT Sponsors

Many FT sponsors have been active in attempting to initiate new FT research. In 1978 USOE awarded contracts to six sponsors, in part to elicit their ideas about potential new FT studies. A review of these and other related reports by FT sponsors yields the following information.

3.1.1 University of Georgia (Ellett, Hawn, Pool and Smock 1979)

This report addresses issues related to a second national FT experiment centering on child outcomes. The authors take the point of view that redoing a large-scale multi-year national study is worthwhile *if* it can be improved through measuring a variety of program outcomes and through gathering implementation and process data that would allow explanations of how model-specific outcomes were achieved. Research to be undertaken through such a national study would address questions such as:

- What is the relative effectiveness of FT models on a common academic performance measure?
- What is the relationship between level of program implementation and outcomes?
- What are the contributions of individual model processes to common learner outcomes? To outcomes unique to each model?

Table 7

FT Research Studies (1977-1980) by Type of Study

Year	Type of Study		
	Planning New FT Research	Improving FT Management and Implementation	New FT Research
1977	• Huron planning study	• Far West Labs implementation paper	
1978	• AAI delayed effects, new models study • Sponsor planning studies	• Oregon implementation study • High/Scope, Kansas joint model study	• Network dissemination study • SDC parent involvement study • Rehab supplemental training study
1979	• Philadelphia delayed effects study • Boon-Young design of new studies	• AMS development of performance indicators • Oregon implementation study	• SDC parent involvement study • Rehab supplemental training study
1980			• National Urban League social implications study

Perhaps the major problem raised in the paper relates to measurement. What sponsor-specific, implementation and process measures should be used in a national evaluation? If such a study were to be funded it would need to be preceded by a period of measure development, perhaps extending the work begun by several FT sponsors (e.g. the Productive Language Achievement Test, the BRACE, several attempts to develop indices of implementation).

3.1.2 Georgia State (Hodges, Carter, Cooper, Burge and Mynatt 1979)

Hodges et al. (1979) work under the premise that the principal vehicle for future FT research will be a large scale multi-year study involving many sponsors and sites. They hope to improve on the national evaluation by offering a complex design that involves implementation of different FT models in the same community, and by broadening the measurement battery to include outcomes (cross-sponsor and sponsor-specific), process, and implementation measures. The types of questions to be addressed through the studies suggested in this report include:

- What is the effect of FT models on classroom practices?
- What happens to classroom practices when the sponsor is removed?
- What is the effect of model implementation on child outcomes?
- Are models differentially effective?
- What is the effect on implementation and outcomes of sponsorship (the typical FT arrangement) vs. knowledge provision (the typical NDN or PIP arrangement)?

3.1.3 University of Kansas (Ramp and Stivers 1979)

Ramp and Stivers suggest that future FT research be focused in two general areas: (1) The relationships between educational processes and products; and (2) sponsors as social change agents. The first focus raises questions about

what philosophies of education work "best", what philosophies of education produce what effects, and what classroom events (across models) produce what child outcomes. A focus on sponsors as change agents raises questions about which sponsor activities ensure the most effective delivery of a program, and what must be known about a community/school prior to implementation of an innovative program. In addition, Ramp and Stivers call for the development of sponsor-specific measures of model implementation. Relevant research questions include:

- What are the relative effects of FT models on common outcome measures?
- What are the effects of individual models on sponsor-specific measures?
- What are the components of an effective model delivery system?
- What are the initial site features that effect implementation?
- What is the relationship between site characteristics and degree of implementation?
- What are the effects of varying the number of adults in the classroom on achievement, program costs, consumer satisfaction?

3.1.4 North Carolina (Olmsted 1979)

The approach of this paper is that a series of small, intensive, well designed studies will yield more and better information than a single large study. Measurement would be intensive and would focus on outcome measures derived from each of FT's components including cognitive and affective development, parent involvement, staff development, and comprehensive services. In addition, measures of model implementation would be crucial. Such small studies would address comparisons both among and within sponsors and would be integrated via meta-analysis techniques (Glass 1977). Several specific studies are proposed to answer questions such as:

- What are the differential effects of FT models on self-concept, locus of control, and achievement?
- What is the relationship between the home environment and child affective and achievement outcomes?
- Are short-term gains in achievement sustained following graduation from FT?
- What is the impact of FT's comprehensive services in terms of degree of utilization, client satisfaction, and child health status?

This paper also proposes that a complementary effort be undertaken to conduct a national survey of FT schools and sponsors, concentrating on documentation of model implementation and local site characteristics. No outcome measurement would be done as this survey would provide descriptive data useful for program management and for the interpretation of results from the smaller evaluations.

3.1.5 University of Oregon (Becker et al. 1979; Carnine 1978)

The point of view offered in these papers is that continuing research should be conducted within FT in order to demonstrate "what can be done to improve problem areas where schools currently fail the most." Demonstration projects are called for in the following areas:

- A large study with improved performance measures, assessment of implementation, and better design.
- A study of the effects of extending FT to grades 4, 5 and 6.
- A study of planned variation in approaches to bilingual education.
- A study of sponsor effectiveness in large cities.

In addition, studies are also recommended to address the following questions:

- What are the effects of FT curricula independent of management systems? For example, does the curriculum (e.g. DISTAR vs. a basal reading program) make a difference when holding the management system (e.g. Direct Instruction sponsorship) constant?

- What are the effects of FT sponsorship? For example, compare the effects of a FT model curriculum as implemented with and without a sponsor.

- What are the effects of different components of the sponsor's management systems in terms of increasing learning opportunities for children?

- What are the effects of various types of parental involvement? What are the effects of models with and without parental education components?

- What are the independent effects of different subject area components on student outcomes?

- What are the effects of different levels of funding on program outcomes?

- What is the impact on fidelity of treatment and student performance of gradually reducing the sponsor's involvement in a site?

- What is the effect of management strength by project officers on contract compliance by sponsors and sites?

- What is the effect of varying academic engaged time?

3.1.6 Southwest Educational Development Laboratory (Lumbley and Kronkosky 1979)

Like some of the other sponsor reports this paper assumes a new, multi-year national FT experiment. It offers alternatives for accomplishing each step of the design and execution of such a study. Questions specific to SEDL's bilingual FT model are posed, e.g.:

- What effect on child outcomes can be expected from differing bilingual program emphases (maintenance vs. transition vs. developmental)?

- How can the transition out of FT be made as smooth as possible?

Further, the paper calls for a study which will allow collection of data that are comparable across sponsors, and also reflective of individual sponsor goals. Finally, a thorough process evaluation is recommended in order to document the sponsor's delivery system and the characteristics of implementation in each

site; provide for formative feedback; and collect implementation and outcome data.

3.1.7 FT Sponsor Evaluation/Research Committee (Gennari 1978)

A task force of sponsor representatives was formed to consider the nature and type of future evaluation/research activities for FT. A preliminary report from this group was issued late in 1978 and made many general recommendations regarding ways in which an appropriate atmosphere could be created for future FT research. For example, the group recommended strongly that future FT research be "multidimensional" (i.e. it should consider all of FT's program components); that the diversity of FT approaches should be respected through sponsor specification of critical and unique characteristics of their programs; that evaluations should focus on the "documentation of program changes as they occur across a variety of educational contexts"; that future research studies be selected and funded on the basis of input from all program stakeholders; and that a program wide ongoing data retrieval system be initiated to facilitate research. In this paper the group did not offer suggestions for specific research/evaluation studies.

3.1.8 Summary of FT Sponsor Research Proposals

After reviewing a few of the sponsor papers some patterns became clear. With one exception, all of the sponsors who completed the planning studies appear to accept the notion of national FT studies which involve, in part, comparisons of sponsor effectiveness on common achievement measures. The sponsors also note quickly that such national studies make sense only in the light of improved research design, measurement of sponsor-specific outcomes, and measurement of implementation and process variables. These sponsors feel that the basic idea

of comparing the effectiveness of alternative educational models makes sense, we just need to do it better and with a more global orientation.

On the other hand, many of the sponsors also call for a series of more diverse studies of smaller scale to address a variety of research questions. Some of the most interesting of these deal with the delayed effects of FT models, the effects of individual components of the FT program, the effects of varying academic learning time, the effects of institutionalization of a sponsor's program (withdrawal of the sponsor), and the effects of sponsorship.

3.2 Review of Research Proposals from Other Planning Efforts

In addition to the FT sponsors, many other planning efforts have been undertaken. Selected efforts not reviewed earlier in this paper will be summarized here. These include papers by the Division of Follow Through, Boon-Young, and the Huron Institute.

3.2.1 Division of Follow Through (1978)

A short paper issued by the Division of FT in 1978 set forth a "blueprint" for studies that could funded in future years. These include research on:

- The impact of noninstructional services.
- Spinoff effects on community, schools, teachers, and parents.
- Delayed effects.
- The effects of preschool experience on FT children.
- The development of new tests and other instrumentation.
- The expansion of FT into grades 4-6.
- The expansion of "small" sponsors that were not part of the national evaluation.
- The relationship between parent involvement and child outcomes.

- The effects of FT on special populations.
- The startup of current sponsors in new sites.
- The startup of new sponsors.
- The utility of resource centers as a vehicle for disseminating research strategies.

Information in several of these areas has already been collected via the Network's study of resource centers, SDC's study of parent involvement, the National Urban League's study of social change, Oregon's implementation study, and AAI's search for new sponsors.

3.2.2 Boon-Young Planning Study (forthcoming)

Motivated in part by the feeling that the national evaluation was too large to be manageable, this planning study was commissioned by the Division of FT in the belief that the potential for acquiring useful knowledge is enhanced by using a small number of sites in each of a number of small scale research efforts. The purposes of the procurement were to provide planning, technical assistance, and design support in three areas:

- o A study/experiment to determine the effects of extending FT beyond grade three.
- o A study/experiment to determine the effects of alternative levels of FT services within and among selected FT models.
- o Planning assistance to determine the capability of LEAs to design, implement, and evaluate their own FT models.

This study is due to be completed in February 1981. Reports should include alternative study designs, feasibility analyses, cost analyses, and so on in each of the three areas.

3.2.3 Huron Institute Planning Study (Garet 1977)

In 1976 the Huron Institute was requested to identify FT evaluation research studies that might be pursued. A set of 25 research areas was identified, and 10

plans were described in some detail. Included were plans for:

- Comparative analyses of sponsor and national evaluation data to determine ways in which sponsor evaluations differed from the national evaluation, factors related to differences, and ways to strengthen FT evaluations.
- Documentation and field testing of sponsor-developed instruments in order to pull together the efforts of sponsors at assessing their specific objectives.
- Establishing criteria for identifying successful models.
- Determining whether appropriate achievement measures (perhaps criterion referenced) are available to measure children's growth over time. Conducting a small scale pilot study.
- Examination of the delivery of comprehensive services in selected FT sites in order to identify effective strategies for service delivery and coordination.
- Exploration and comparison of the forms of parent involvement in FT in an effort to assess the efficacy of various strategies.
- Documentation and exploration of the process of model development.
- Investigation of the relationship between "contact time" and achievement.
- Exploration of explanations for within-sponsor site to site variation in child achievement.
- Assessment of the feasibility of using "site inspectors", e.g. site visitors, interviewers, participant-observers, in an evaluation of FT.

4.0 Summary and Recommendations

The thesis of this paper is that there is more to be learned from FT in its current framework. Three ways of generating new knowledge from FT have been discussed: (1) secondary analysis of extant data; (2) compilation and analysis of follow-up data; and (3) completely new studies. Although some of the ideas reviewed in this paper may be infeasible or impractical, a case has been made that there are many areas in which we can learn from FT, and in which FT can demonstrate effective ways of educating young children.

While many of the research questions/proposals reviewed in this paper have merit, two are recommended in this concluding section for special consideration. The first of these is research on the potential delayed effects of FT. There is an abundance of reasons to invest resources in this area. First, the long-term effects of educational programs is an important area of research. Of what purpose is education if there is no prospect for lasting effects? Second, there are ample precedents for studying delayed effects (e.g. Lazar et al. 1977); precedents which have had enormous impact on Head Start. Third, impact on "life chances" variables is the ultimate goal of the FT program and should be investigated if at all possible. Fourth, many FT sponsors hypothesize differential patterns of long-term effects and are in favor of such a study. Fifth, the data to perform delayed effects studies exist in a reasonable (though not completely adequate) form. Finally, since studies of post-FT effects would be secondary analytic in nature they would be relatively inexpensive and could provide information in a timely manner. Results from an investigation into the delayed effects of FT on children could well have important implications for the design and funding of FT as well as other early childhood education programs.

Delayed effects studies could well be expanded to address the long-term effects of FT on parents and teachers. This is an area of research that has received much less attention than studies of delayed effects on children. Examples of relevant questions that would entail new data collection include:

- What are the long-term perceptions of children and parents about FT? Do they remember FT? If so, in what respects?
- What are the impacts of FT on parent level life chances variables (e.g. continued education, employment)? Did FT lead parents to further their education? To gain new employment skills?
- What are the impacts of FT on teacher level life chances variables? Did FT provide useful training for teachers? Did it further their employment status?

A second area of research also merits further discussion. Some of the sponsor research papers venture the notion that we should study sponsorship itself. This is an important idea that could lead to the integration of FT studies with studies of other mechanisms for change in schooling. The concept and operationalization of sponsorship is perhaps FT's most creative contribution to educational change. Yet, FT research has been so focused on child outcomes that very little effort has been devoted to understanding the key concept of sponsorship, defining it, or comparing it with other mechanisms for change.

An investigation of sponsorship might begin with a comparative analysis of sponsorship vs. other change mechanisms (e.g. the National Diffusion Network, or Project Information Packages). While we might equate PIPs with the simple provision of information, and the NDN with the provision of information plus limited technical assistance, sponsorship entails (at the very least) the concept of a long-term relationship between the sponsor and an LEA as well as the introduction of a new curriculum or use of different teaching methods. The fact that a FT sponsor is committed to his/her model, and remains in a site for many years to help implement the model is a radical departure from other change strategies.

Of key interest to a preliminary descriptive analysis would be questions such as: What assumptions about the educational change process do different change mechanisms make? What are the differences in type of services provided? In the intensity of services provided? Given such a descriptive analysis it would be possible to proceed with selected studies of the effects of different change mechanisms. For example, what is the effect of a curriculum as implemented under FT sponsorship vs. under some other change mechanism? What does the curriculum look like under each implementation strategy? What are the

respective outcomes? On the other hand, one can look within FT and ask, what is the effect of a given curriculum when implemented by different FT sponsors? Again, what does the implemented curriculum look like? What are the outcomes? Finally, it would be important to consider what happens to a program when sponsorship is withdrawn? Do programs which are implemented via sponsorship stay implemented longer than those implemented by some other change mechanism?

In closing it must be stated that all the planning done to date, including the effort put into this and the other papers prepared for NIE, will go for naught unless some of the organizational and structural constraints that have hampered FT research for the past few years are removed. This does not mean that the program has to receive a guarantee of funding--research can be planned under conditions of uncertainty. It is critical, however, to find a way to avoid the situation of the recent past in which the Division of Follow Through and the Office of Program Evaluation were often at odds about the appropriate directions to take with respect to FT research. This could be accomplished by assigning responsibility to a single group that would have the authority to plan and execute FT research and evaluation, or by dividing responsibility among multiple groups each having their own budget (as currently appears to be the case with the Division of Follow Through and NIE having separate agendas and budgets). The problem to avoid is that of having multiple groups involved, each with veto power over the other's activities. If this continues to be the case the question will not be whether there is more to be learned from FT, but whether the constraints placed on the planning process will allow anything to be learned.

References

Becker, W.C. and S. Engelmann. Analysis of achievement data on six cohorts of low income children from 20 school districts in the University of Oregon Direct Instruction Follow Through Model, Technical report 78-1. Eugene, OR: University of Oregon Follow Through Project, December 1978.

Becker, W.C. et al. Planning information for Follow Through experiments. Eugene, OR: Direct Instruction Follow Through Model, University of Oregon, 1979.

Bereiter, C. and M. Kurland. "A constructive look at Follow Through results." Unpublished manuscript. Toronto, Ontario, Canada: Ontario Institute for Studies in Education, October 1978.

Bock, G., L. Stebbins and E. Proper. Education as experimentation: A planned variation model. Volume IV-B: Effects of Follow Through models. Report to the U. S. Office of Education pursuant to Contract No. 300-75-0134. Cambridge, MA: Abt Associates Inc., 1977.

Boruch R. and J. Reis. "The student, evaluative data, and secondary analysis." In L. Sechrest (Ed.) New directions for program evaluation: Training program evaluators. San Francisco, CA: Jossey-Bass Inc., 1980.

Camilli, G. A reanalysis of the effect of Follow Through on cognitive and affective development. Unpublished Ph.D. thesis. Boulder, CO: University of Colorado, Laboratory of Educational Research, 1980.

Campbell, D.T. "Reforms as experiments." American psychologist, 1969, 24, (4), 409-429.

Campbell, D.T. "Methods for the experimenting society." Address presented to the American Psychological Association. Washington, DC: 1971.

Carnine, D. "Possible research topics for FT." Unpublished memorandum to policy makers involved in planning for Follow Through. Eugene, OR: Direct Instruction Follow Through Model, University of Oregon, 1978.

Cloud, K., R.K. Rentfrow and J. Hildebrandt. "Later school success of Follow Through graduates in Lakewood, New Jersey Public Schools." Tucson, AZ: Center for Educational Research and Development, University of Arizona, July 1979.

Cook, T.D. "The potential and limitations of secondary evaluations." In M.W. Apple, M.J. Subkoviak and H.S. Lufler (Eds.) Educational evaluation: Analysis and responsibility. Berkeley, CA: McCutchan Publishing Company, 1974.

Division of Follow Through. "Blueprint for studies to be funded in Follow Through." Washington, DC: Division of Follow Through, U. S. Department of Education, September 1978.

Edwards, R. and J. Bridewell. "Following through on Follow Through." Paper presented at the annual meeting of the American Educational Research Association. San Francisco, CA: April 1979.

Ellett, C., H. Hawn, K. Pool, and C. Smock. Planning information study for future Follow Through experiments. Athens, GA: Mathemagenic Activities Program, University of Georgia, 1979.

Garet, M. Follow Through evaluation studies in the next two years: Ten potential plans. Cambridge, MA: Huron Institute, 1977.

Gennari, T. "A summary of discussions by the evaluation/research committee." Unpublished memorandum to Follow Through Forward Planning Task Force. Pittsburgh, PA: Learning Research and Development Center, University of Pittsburgh, December 1978.

Glass, G.V. "Integrating findings: The meta-analysis of research." In L.S. Shulman (Ed.) Review of research in education, Volume 5. Itasca, IL: F.E. Peacock Publishers, Inc. 1977.

Goodrich, R.L. and R.G. St.Pierre. Opportunities for studying later effects of Follow Through. Report prepared for the U.S. Office of Education pursuant to Contract No. 300-78-0443. Cambridge, MA: Abt Associates Inc., December 1979.

Haney, W. and N. Pennington. Reanalysis of Follow Through parent and teacher data from spring 1975. Report to the U.S. Office of Education pursuant to Contract No. OEC-74-0394. Cambridge, MA: Huron Institute, 1978.

Hodges, W.L. and R. Sheehan. "Ten years of evaluation efforts: The work of Follow Through sponsors." Paper presented at the annual meeting of the American Educational Research Association. Toronto, Ontario, Canada, 1978.

Hodges, W., H. Carter, M. Cooper, A. Burge, and G. Mynatt. Issues in planning for Follow Through experiments: A multiple evaluation strategy. Atlanta, GA: College of Education, Georgia State University, 1979.

House, E.R., G.V. Glass, L.D. McLean and D.F. Walker. "No simple answer: Critique of the Follow Through evaluation." Harvard educational review, 1978, 48, (2), 128-160.

Lazar, I., V.R. Hubbell, H. Murray, M. Rosche and J. Royce. Summary: The persistence of preschool effects. Report to the Administration for Children, Youth and Families, pursuant to Grant No. 18-76-07843. Ithaca, NY: Cornell University, 1977.

Lumbley, J. and P. Kronkosky. Planning information for Follow Through experiments. Austin, TX: Follow Through Program, Southwest Educational Development Laboratory, 1979.

Madaus, G.F. and T. Marx. Report on the Follow Through later grades study. Cambridge, MA: Huron Institute, April 1973.

Middleton, T. and L.A. Durgin. Summary of Follow Through data. Report to the U.S. Office of Education pursuant to Contract Nos. OEC-0-522480-4633 (100) and OEC-0-0582. Menlo Park, CA: SRI International, 1978.

Olmsted, P. Planning information for Follow Through experiments. Chapel Hill, NC: Parent Education Model, University of North Carolina, 1979.

Ramp, G. and M. Stivers. Planning information for Follow Through experiments. Lawrence, KS: Department of Human Development, University of Kansas, 1979.

Riley, G.R. Follow Through program evaluation, 1976-77. Wichita, KS: Wichita Public Schools, Division of Research, Planning and Development Services, January 1978.

School District of Philadelphia. The long-term effects of seven Follow Through models in Philadelphia. Research proposal submitted to the U. S. Office of Education. Philadelphia, PA: The School District of Philadelphia, 1980.

Seitz, V., N. Apfel, and C. Efron. "Long-term effects of early intervention: The New Haven Project." Paper presented at the annual meeting of the American Association for the Advancement of Science. Denver, CO: 1977.

Shive, G., P. Meleen, S. Harris, L. Vaughn, and M. Grogan. A policy study of resource centers for the Division of Follow Through. Andover, MA: The Network, January 1980.

Stebbins, L.B., R.G. St.Pierre, E.C. Proper, R.B. Anderson and T.R. Cerva. Education as experimentation: A planned variation model. Volume IV-A: An evaluation of Follow Through. Report to the U.S. Office of Education pursuant to Contract No. 300-75-0134. Cambridge, MA: Abt Associates Inc., 1977.

Weber, B. and M. Fuhrmann. "A study of district A's former Follow Through students' retention of basic skills after six years out of the program." Paper prepared for District A, 1978.

Wholey, J. "New directions for Follow Through." Washington, DC: U. S. Department of Education, December 1979.

Wolff, M. and A. Stein. Six months later, Head Start evaluation project. New York, NY: Yeshiva University, Ferkauf Graduate School of Education, 1966.